



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/877,035	06/11/2001	Toshihiko Munetsugu	P21107	9810
7055	7590	06/25/2007	EXAMINER	
GREENBLUM & BERNSTEIN, P.L.C. 1950 ROLAND CLARKE PLACE RESTON, VA 20191			TRAN. QUOC A	
ART UNIT		PAPER NUMBER		
2176				
NOTIFICATION DATE		DELIVERY MODE		
06/25/2007		ELECTRONIC		

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

gbpatent@gbpatent.com
pto@gbpatent.com



UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

MAILED

JUN 25 2007

Technology Center 2100

**BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES**

Application Number: 09/877,035

Filing Date: June 11, 2001

Appellant(s): MUNETSUGU ET AL.

Bruce H. Bernstein
For Appellant

EXAMINER'S ANSWER

This is in response to the appeal brief filed 02-27-2007 appealing from the Office action mailed 08-24-2005.

(1) Real Party in Interest

A statement identifying by name the real party in interest is contained in the brief.

(2) Related Appeals and Interferences

The examiner is not aware of any related appeals, interferences, or judicial proceedings, which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

(3) Status of Claims

The statement of the status of claims contained in the brief is correct.

(4) Status of Amendments After Final

The appellant's statement of the status of amendments after final rejection contained in the brief is correct.

(5) Summary of Claimed Subject Matter

The summary of claimed subject matter contained in the brief is correct.

(6) Grounds of Rejection to be Reviewed on Appeal

The appellant's statement of the grounds of rejection to be reviewed on appeal is correct.

(7) Claims Appendix

The copy of the appealed claims contained in the Appendix to the brief is correct.

(8) Evidence Relied Upon

US005969716A	Davis et al.	filed 08/06/1996
US006560640B2	Jain et al.	filed 08/14/1998

(9) Grounds of Rejection

The following ground(s) of rejection are applicable to the appealed claims:

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 1-4, 11-13 and 21-27 are rejected under 35 U.S.C. 103(a) as being unpatentable over Davis et al. US005969716A - filed Aug. 6, 1996 (hereinafter Davis '697), in view of Jain et al. US006360234B2 - filed Aug. 14, 1998 (hereinafter Jain'234).

In regard to independent claim 1, analyzer that receives as input structure description data in which media content is described (Davis '697 at col. 2, lines 50-55, discloses an automatic time-based media processing system, wherein the media signal is processed in a media parser to obtain descriptive representation of its content), **the media content being continuous audiovisual information** (Davis '697 at col. 2, lines 53-67, discloses the representation of the media signal, such as its prosody (i.e., its pitch pattern), or in the case of music, its chord structures), **the structure description data describing types of media included in the media content**, (Davis '697 at col. 2, lines 31-67, discloses a media parser to obtain descriptive representations of its contents. Each content representation is data that

provides information about the media signal, and is functionally dependent on the media signal, such as: frames, timecodes, movies, television programs, music videos, etc.), **and a plurality of segments that use the media. Expressed in time information, wherein the analyzer extracts the time information of the segments from the structure description data** (Davis '697 at col. 11, line 38 through col. 12, line 5, also see Fig. 8, discloses a media parser to obtain descriptive representations of its contents, wherein FIG. 8 illustrates a graphical user interface that can be presented on the screen of the display (item 28), this user interface consists of a number of different sections, which are arranged in columnar form, such as: the media signals, and content representations of them, a timeline format, a ruler (item 36), that depicts increments of time, e.g. seconds or any suitable metric can be represented by the ruler, for example the indices of the events in a sequential), **a converter that automatically organizes the types of media** (Davis '697 at col. 11, line 38 through col. 12, line 5, also see Fig. 8, discloses a media parser to obtain descriptive representations of its contents, wherein FIG. 8 illustrates a graphical user interface that can be presented on the screen of the display (item 28), this user interface consists of a number of different sections, which are arranged in columnar form, such as: the media signals, and content representations of them, a timeline format, a ruler (item 36), that depicts increments of time, e.g. seconds or any suitable metric can be represented by the ruler, for example the indices of the events in a sequential), **automatically arranges the types of media** (Davis '697 at col. 2, lines 31-67, discloses an automatic time-based media processing system, wherein the media signal is processed in a media parser to obtain descriptive representation of its content. Each content representation is data that provides information about the media signal, and is

Art Unit: 2176

functionally dependent on the media signal, such as: frames, timecodes, movies, television programs, music videos, etc.),

Davis '697 does not explicitly teach, **addresses indicating locations of the media content, and the addresses per extracted time information, and addresses in an order of representation thereby automatically convert the structure description data into representation description data that specifies an order of representation and synchronization information of the segments**, however (Jain'234 at col. 1, line 1 through col. 2, line 15, discloses a multimedia cataloger with plurality sync encoders, that is automatically watch, listen to and read a video stream, the multimedia cataloger intelligently extracts metadata-key-frames, time codes, textual information and an audio profile from the video in real-time, wherein frame-accurate index that provides immediate, non-linear access to any segment of the media, further more detailing Jain'234 at col. 5, line 6 through col. 7, line 20, discloses the detailing of "Vidsync" process of encoding MPEG files, wherein GUI is utilizing to mark in- and out-times, and type in associated alphanumeric data. Each bar in the Clip Track consists of a user-defined group of metadata fields that are application specific. The bar length is timespan from intime to outtime. Clips may be overlapping and metadata may include: Story Title, Report, Location, Shot Date, Air Date, Keywords, Summary, and so on).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Davis '697 teaching, that provides an automatic time-based media processing system, wherein the media signal is processed in a media parser to obtain descriptive representation of its content, to include a means of indicating the address locations of the media content, and the addresses per extracted time information, and addresses

in an order of representation thereby automatically convert the structure description data into representation description data that specifies an order of representation and synchronization information of the segments of Jain'234. One of ordinary skill in the art would have been motivated to modify this combination to improve the time-based media processing system, which is capable of providing high quality, adaptive, efficient, re-usable of media content without the without requiring a significant level of skill on the part of the user, and is therefore suited for use by the average consumer (as taught by Davis'716 at col. 2, lines 7-18).

In regard to independent claim 11, incorporate substantially similar subject matter as cited in claim 1 above, and further view of the following, and is similarly rejected along the same rationale, ...**selection condition... media content score**, however (Jain'234 at col. 5, line 6 through col. 7, line 20, discloses the detailing of "Vidsync" process of encoding MPEG files, wherein GUI is utilizing to mark in- and out-times, and type in associated alphanumeric data. Each bar in the Clip Track consists of a user-defined group of metadata fields that are application specific. The bar length is timespan from intime to outtime. Clips may be overlapping and metadata may include: Story Title, Report, Location, Shot Date, Air Date, Keywords, Summary, and so on), Examiner read the above in the broadest reasonable interpretation to the claim limitation, wherein media content score would have been an obvious variant of media meta data and selection condition would have been an obvious variant of GUI is utilizing to mark...to a person of ordinary skill in the art at the time the invention was made.

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Davis '697 teaching, that provides an automatic time-based media processing system, wherein the media signal is processed in a media parser to

obtain descriptive representation of its content, to include a means of indicating the address locations of the media content, and the addresses per extracted time information, and addresses in an order of representation thereby automatically convert the structure description data into representation description data that specifies an order of representation and synchronization information of the segments of Jain'234. One of ordinary skill in the art would have been motivated to modify this combination to improve the time-based media processing system, which is capable of providing high quality, adaptive, efficient, re-usable of media content without the without requiring a significant level of skill on the part of the user, and is therefore suited for use by the average consumer (as taught by Davis'716 at col. 2, lines 7-18).

In regard to dependent claims 2 and 4, incorporate substantially similar subject matter as cited in claim 1 above, and are similarly rejected along the same rationale.

In regard to dependent claim 3, wherein the representation decryption data is a SMIL document however (Jain'234 at col. 8 lines 32-49, discloses media format can be in SMIL).

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Davis '697 teaching, that provides an automatic time-based media processing system, wherein the media signal is processed in a media parser to obtain descriptive representation of its content, to include a means of representation decryption data is a SMIL document of Jain'234. One of ordinary skill in the art would have been motivated to modify this combination to improve the time-based media processing system, which is capable of providing high quality, adaptive, efficient, re-usable of media content without the without

requiring a significant level of skill on the part of the user, and is therefore suited for use by the average consumer (as taught by Davis'716 at col. 2, lines 7-18).

In regard to dependent claim 12, incorporate substantially similar subject matter as cited in claim 11 above, and further view of the following, and is similarly rejected along the same rationale, ...**a network that connects said server and said client ...**, however (Jain'234 at col. 5, lines 23-50, also see Fig. 4, provides a data network (item 250) environment, wherein all machines are connected using standardized TCP/IP network protocol).

In regard to dependent claim 13, incorporate substantially similar subject matter as cited in claims 11 and 12 above, and is similarly rejected along the same rationale.

In regard to dependent claims 22, 23, 26 and 27, incorporate substantially similar subject matter as cited in claim 1 above, and are similarly rejected along the same rationale.

In regard to dependent claims 21 and 25, incorporate substantially similar subject matter as cited in claim 11 above, and are similarly rejected along the same rationale.

In regard to dependent claim 24, incorporate substantially similar subject matter as cited in claim 3 above, and is similarly rejected along the same rationale.

(10) Response to Argument

Brief summary of prior art of records:

Davis discloses a media parser to obtain descriptive representation of media content (i.e. time-based, data types, types) to manipulate the representation of the content of the media (i.e. sync, substitute, temporal compression, dilation and parametric special effect).

Art Unit: 2176

Jain discloses a web video streaming as Meta data capture and out put by format such as XML, SMIL described by W3C.

Response to Arguments:

Beginning on page 7 of the appeal brief (hereinafter the brief), Appellant argues the following issues, which are accordingly addressed below.

Appellant argues on pages 7-10 of the brief that the Rejection of Claims 12, 13, 21 and 26 under 35 U.S.C. 112, 1st and 2nd paragraph are improper, and should be reversed.

The examiner respectfully agrees. The Examiner respectfully has withdrawn the rejections of 35 U.S.C. 112, 1st and 2nd (see the brief page 6 bottom).

Appellant argues on pages 10-14 of the brief that the resulting of Davis and Jain in combination do not teach, **the structure description data describing types of media included in the media content, and a plurality of segments that use the media. Expressed in time information, wherein the analyzer extracts the time information of the segments from the structure description data automatically arranges the types of media addresses indicating locations of the media content, and the addresses per extracted time information, and addresses in an order of representation thereby automatically convert the structure description data into representation description data that specifies an order of representation and synchronization information of the segments of claim 1, and not proper to combine Davis and Jain** (the same arguments are substantially repeated for independent claim 11, dependent claims 2-4, 12-13, 21-27 pending).

The examiner respectfully disagrees. The examiner respectfully notes that (see Davis at col. 2, lines 20-55), discloses Content Representation, automatically, semi-automatically, and manually generated descriptive data that represent the content of media signals, functional relationships that operate on content representations and media signals to compute new media content, media parser for time-based media processing systems, which manipulate representations of media content in order to compute new media content. The invention is intended to support a paradigm shift from the direct manipulation of simple temporal representations of media (frames, timecodes, etc.), to the interactive computation of new media from higher-level representations of media content and functional dependencies among them. This paradigm of media processing and composition enables the production of traditional media (e.g., movies, television programs, music videos, etc.) and a plurality of segments that use the media.

For further more support the above, the Examiner introduces Jain reference at col. 7, lines 1-10, also see Fig. 7 also see Table 1 col. 7 lines 30-60, discloses Metadata Index Object Model, the main object is The main object, the Metadata Track Index Manager 402, is the manager of the entire index of metadata. It is extensible in that it allows registration of individual metadata track data types, and then manages the commitment of instances of that data into the index by feature extractors. There is one global metadata structure (the Session Level metadata 404) that is not time based, and contains metadata that pertains to the entire video. Here, for example, is where the information for managing and time-synching the encoded video resides (digital video ID's and actual start time offsets). User defined annotations may also exist here. Each of the metadata tracks is a collection of data objects 406, 408, 410, 412, etc. that hold the metadata for

a specific feature extractor, and are sequenced in time according to their in- and out-times, and metadata index also provides access for outputting metadata (data read-out) used by the Output Filters.

Also (see Jain at col. 8 lines 35-50) discloses the metadata the metadata may be output in a variety of formats such as Virage Data Format (VDF) 562, HTML 564, XML 566, SMIL 568 and other 570, which are managed by the Output Filter Manager 560. A VDF API and Toolkit may be licensed from Virage of San Mateo, California. Furthermore, the use of the format is described in "Virage VDF Toolkit Programmer's Reference". One reference for the extensible Mark-up Language (XML) is the following URL: <http://www.w3.org/TR/REC-xml>, which is a subpage for the W3C. Also, information on Synchronized Multimedia Integration Language (SMIL) may be accessed at the W3C site.

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Jain's result (i.e. the metadata may be output in a variety of formats such as (VDF), HTML, XML, SMIL and other) into Davis's teaching, that provides an automatic time-based media processing system, wherein the media signal is processed in a media parser to obtain descriptive representation of its content.

One of ordinary skill in the art would have been motivated to modify this combination, because they are from the same field of endeavor of video, audio authoring, in the time-based media processing system, which is capable of providing high quality, adaptive, efficient, reusable of media content without the requiring a significant level of skill on the part of the user, and is therefore suited for use by the average consumer (as taught by Davis'716 at col. 2, lines 7-18).

Appellant argues on pages 14-26 of the brief that the resulting of Davis and Jain in combination do not teach, some of the claimed limitation of claims 2-4, 12-13, and 21-27, which (the same arguments are substantially repeated here from response to the brief for independent claim 11, dependent claims 2-4, 12-13, 21-27 cited above), and further view of the following:

"describes a set of alternative data to the media content", **"wherein the representation description data is a SMIL document"**, **"disclose or suggest a converter that "describes, in the representation description data, selection conditions for selecting the media content and alternative data"**, **"selects and represents one of the media content and the alternative data in accordance with the selection conditions"**, **"selection condition" "media content score"** , **"addresses indicating locations of the media content"**, **"the selector and converter"**, **"a client comprising the presenter"**, **"a network that connects said server and said client,** **wherein the representation description data is communicated between said server and said client"**.

The examiner respectfully disagrees. The examiner respectfully notes that (see Davis at col. 2, lines 20-55), discloses Content Representation, automatically, semi-automatically, and manually generated descriptive data that represent the content of media signals, functional relationships that operate on content representations and media signals to compute new media content, media parser for time-based media processing systems, which manipulate representations of media content in order to compute new media content. The invention is intended to support a paradigm shift from the direct manipulation of simple temporal representations of media (frames, timecodes, etc.), to the interactive computation of new media from higher-level representations of media content and functional dependencies among them.

This paradigm of media processing and composition enables the production of traditional media (e.g., movies, television programs, music videos, etc.) and a plurality of segments that use the media.

For further more support the above, the Examiner introduces Jain reference at col. 7, lines 1-10, also see Fig. 7 also see Table 1 col. 7 lines 30-60, discloses Metadata Index Object Model, the main object is The main object, the Metadata Track Index Manager 402, is the manager of the entire index of metadata. It is extensible in that it allows registration of individual metadata track data types, and then manages the commitment of instances of that data into the index by feature extractors. There is one global metadata structure (the Session Level metadata 404) that is not time based, and contains metadata that pertains to the entire video. Here, for example, is where the information for managing and time-synching the encoded video resides (digital video ID's and actual start time offsets). User defined annotations may also exist here. Each of the metadata tracks is a collection of data objects 406, 408, 410, 412, etc. that hold the metadata for a specific feature extractor, and are sequenced in time according to their in- and out-times, and metadata index also provides access for outputting metadata (data read-out) used by the Output Filters.

Also (see Jain at col. 8 lines 35-50) discloses the metadata the metadata may be output in a variety of formats such as Virage Data Format (VDF) 562, HTML 564, XML 566, SMIL 568 and other 570, which are managed by the Output Filter Manager 560. A VDF API and Toolkit may be licensed from Virage of San Mateo, California. Furthermore, the use of the format is described in "Virage VDF Toolkit Programmer's Reference". One reference for the extensible Mark-up Language (XML) is the following URL: <http://www.w3.org/TR/REC-xml>, which is a

Art Unit: 2176

subpage for the W3C. Also, information on Synchronized Multimedia Integration Language (SMIL) may be accessed at the W3C site.

It would have been obvious to a person of ordinary skill in the art at the time the invention was made to have modified Jain's result (i.e. the metadata may be output in a variety of formats such as (VDF), HTML, XML, SMIL and other and "describes a set of alternative data to the media content", "wherein the representation description data is a SMIL document", "disclose or suggest a converter that "describes, in the representation description data, selection conditions for selecting the media content and alternative data", "selects and represents one of the media content and the alternative data in accordance with the selection conditions", "selection condition" "media content score" , "addresses indicating locations of the media content", "the selector and converter", " a client comprising the presenter", "a network that connects said server and said client, wherein the representation description data is communicated between said server and said client") into Davis's teaching, that provides an automatic time-based media processing system, wherein the media signal is processed in a media parser to obtain descriptive representation of its content.

One of ordinary skill in the art would have been motivated to modify this combination, because they are from the same field of endeavor of video, audio authoring, in the time-based media processing system, which is capable of providing high quality, adaptive, efficient, reusable of media content without the requiring a significant level of skill on the part of the user, and is therefore suited for use by the average consumer (as taught by Davis'716 at col. 2, lines 7-18).

Therefor the Examiner respectfully maintains the rejection of independent claims 1, 11, dependent claims 2-4, 12-13, and 21-27 and should be sustained.

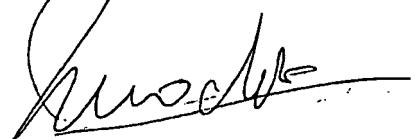
(11) Related Proceeding(s) Appendix

No decision rendered by a court or the Board is identified by the examiner in the Related Appeals and Interferences section of this examiner's answer.

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,

Quoc A. Tran



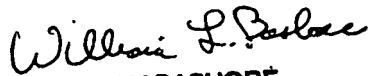
Conferees:

Heather R. Herndon



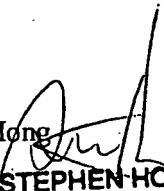
Heather R.
HEATHER R. HERNDON
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100

William L. Bashore



William L. Bashore
WILLIAM BASHORE
PRIMARY EXAMINER

Stephen S. Hong



STEPHEN HONG
SUPERVISORY PATENT EXAMINER